

Pu Li and Quoc Dong Vu

Identification of parameter correlations for parameter estimation in dynamic biological models

Original published in:

BMC systems biology. - London : BioMed Central. – 7 (2013), 91 (22 September 2013), pp 1–12.

DOI: 10.1186/1752-0509-7-91

ISSN (online): 1752-0509

URL: <http://www.biomedcentral.com/1752-0509/7/91>

[Visited: 2015-01-26]



This work is licensed under a
[Creative Commons Attribution 2.0 Generic License](http://creativecommons.org/licenses/by/2.0/).
[<http://creativecommons.org/licenses/by/2.0/>]

METHODOLOGY ARTICLE

Open Access

Identification of parameter correlations for parameter estimation in dynamic biological models

Pu Li* and Quoc Dong Vu

Abstract

Background: One of the challenging tasks in systems biology is parameter estimation in nonlinear dynamic models. A biological model usually contains a large number of correlated parameters leading to non-identifiability problems. Although many approaches have been developed to address both structural and practical non-identifiability problems, very few studies have been made to systematically investigate parameter correlations.

Results: In this study we present an approach that is able to identify both pairwise parameter correlations and higher order interrelationships among parameters in nonlinear dynamic models. Correlations are interpreted as surfaces in the subspaces of correlated parameters. Based on the correlation information obtained in this way both structural and practical non-identifiability can be clarified. Moreover, it can be concluded from the correlation analysis that a minimum number of data sets with different inputs for experimental design are needed to relieve the parameter correlations, which corresponds to the maximum number of correlated parameters among the correlation groups.

Conclusions: The information of pairwise and higher order interrelationships among parameters in biological models gives a deeper insight into the cause of non-identifiability problems. The result of our correlation analysis provides a necessary condition for experimental design in order to acquire suitable measurement data for unique parameter estimation.

Keywords: Parameter estimation, Identifiability, Parameter correlation, Data sets with different inputs, Zero residual surfaces, Experimental design

Background

Parameter estimation of dynamic biological models described by nonlinear ordinary differential equations (ODEs) poses a critical challenge. A special feature of biological models is that they usually contain a large number of parameters among which correlations exist [1,2]. In general, the quality of estimation results depends on the quality of data acquisition, the quality of the fitting method, and the quality of the model. Good experiment design can lead to highly informative data which will enhance the accuracy and identifiability of model parameters. Therefore, the task of parameter

estimation demands an interactive endeavour of experiment and computation [3,4].

To fit parameters to measured data a numerical method for solving an optimization problem is required. Available methods for carrying out this task can be classified into deterministic approaches (e.g., multiple shooting [5,6], collocation on finite elements [7], global approaches [8,9]) and stochastic approaches (e.g. simulated annealing [10], genetic algorithms [11], and scatter search [12]). Using these approaches, model parameters can be well fitted to measured time courses provided from either experiment (*in vivo*) or simulation (*in silico*), i.e. high quality fits with minimal residual values can be obtained by global optimization methods.

However, even such a good fitting cannot guarantee unique parameter estimation, due to correlations among the parameters. The correlation phenomenon can be

* Correspondence: pu.li@tu-ilmenau.de
Department of Simulation and Optimal Processes, Institute of Automation and Systems Engineering, Ilmenau University of Technology, P. O. Box 100565, 98684 Ilmenau, Germany

explained by the biological background, e.g. genes or proteins which are expressed in a correlated manner, correlations of expression levels between cells. As a consequence, certain regions in the parameter space correspond to good model predictions. It means that the residual value (quadratic error) remains low even if the parameters vary in certain regions. By testing 17 biological models, Gutenkunst et al. [13] concluded that collective fits to large amounts of ideal time-series data lead to the fact that some eigenvectors are orders of magnitudes better constrained than others.

Correlated parameters are non-identifiable. If the non-identifiability does not change for any data, these parameters are called structurally non-identifiable. On the contrary, if the non-identifiability can be remedied by data improvement, they are practically non-identifiable [14,15]. Identifiability analysis represents an important ongoing topic in the literature which can be in general categorized into two major groups: *a priori* and *a posteriori* methods [1,16]. Without any requirement of measurement data, global (structural) identifiability can be determined by *a priori* methods [17-19]. Since these methods are normally based on differential algebra, their application to high dimensional complex models can be limited.

The *a posteriori* methods reveal practical identifiability properties based on results from fitting parameters to available data sets. In most studies, correlations are detected by analysing the sensitivity matrix and the Fisher information matrix (FIM) [1,16,20-23], from which local confidence regions of parameter solutions can be obtained. Sensitivity analysis is well suitable to linear models but will have limitations for highly nonlinear models [14,24].

Recently, Raue et al. [15] proposed to use profile likelihood to detect non-identifiability for partially observable models. The parameter space is explored for each parameter by repeatedly fitting the model to a given data set, which then provides a likelihood-based confidence region for each parameter. Results from this method show that the number of practically non-identifiable parameters will decrease when more data sets are used [25].

An aim of identifiability analysis is to determine if the parameters of a model are identifiable or not, i.e. whether its parameters can be uniquely estimated. The profile likelihood approach can also offer information on the correlated relations among the parameters [15,25-27]. The information on parameter correlations (e.g. correlated groups, correlated forms in a group etc.) is important for experimental design, so that a series of experimental runs with determined conditions can be carried out to acquire proper measurement data sets for improving the quality of parameter estimation.

Very few studies have been made to investigate parameter correlations in biological models. Yao et al. [21]

used the rank of the sensitivity matrix to determine the number of estimable parameters. However, the subsets of correlated parameters cannot be identified based on this result. Chu and Hahn [28] proposed to check the parallel columns in the sensitivity matrix to determine parameter subsets in which the parameters are pairwise correlated. Quaiser and Mönnigmann [29] proposed a method to rank the parameters from least estimable to most estimable. These methods, however, cannot identify parameter groups in which more than two parameters are correlated together but not in pairwise, i.e. the corresponding columns in the sensitivity matrix are linearly dependent but not parallel. Such correlations are called higher order interrelationships among parameters [16].

In this paper, "parameter correlations" means a group of parameters in the model equations which are mathematically related to each other through some implicit functions, i.e. among the parameters there is a functional relationship [15,26,27]. Correlated parameters will be structurally non-identifiable, if the functional relationship does not depend on the control variables which determine experimental conditions and thus measured data. On the other hand, they will be practically non-identifiable, if the functional relationship depends on the control variables.

In this paper, we present an approach which is able to identify both pairwise and higher order parameter correlations. Our approach is based on analysis of linear dependences of the first order partial derivative functions of model equations. In a given model there may be a number of groups with different number of correlated parameters. We propose to identify these groups by analysing the correlations of the columns of the state sensitivity matrix which can be derived directly from the right-hand side of the ODEs. Therefore, the method proposed in this paper is *a priori* in nature, which means that the parameter correlations considered in this paper are not from the results of data-based estimation. A geometric interpretation of parameter correlations is also presented. Using this approach, groups of correlated parameters and the types of correlations can be identified and, hence, the parameter identifiability issue can be addressed. Moreover, the relationship between parameter correlations and the control inputs can be derived. As a result, both structural and practical non-identifiabilities can be identified by the proposed approach.

In the case of practical non-identifiability, the parameter correlations can be relieved by specifying the values of control inputs for experimental design. Based on the correlation analysis, the maximum number of parameters among the correlation groups can be determined, which corresponds to the minimum number of data sets with different inputs required for uniquely estimating

the parameters of the model under consideration. Numerical results of parameter estimation of a three-step-pathway model clearly demonstrate the efficacy of the proposed approach.

Methods

Identification of parameter correlations

We consider nonlinear model systems described by

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t), \mathbf{p}) \quad (1)$$

$$\mathbf{y}(t) = \mathbf{h}(\mathbf{x}(t), \mathbf{u}(t), \mathbf{q}) \quad (2)$$

where $\mathbf{x}(t) \in R^n$ is the state vector, $\mathbf{u}(t) \in R^m$ the control vector, and $\mathbf{y}(t) \in R^r$ the output vector, respectively. In this study, two different sets of parameters, i.e. $\mathbf{p} \in R^{NP}$ in the state equations and $\mathbf{q} \in R^{NQ}$ in the output equations, are considered. In most cases the number of parameters in the state equations is much larger than that in the output equations. Since the correlations of the parameters in the output Equation (2) are easier to identify, we concentrate on the analysis and identification of correlations of the parameters in the state Equation (1).

The equation of the state sensitivity matrix can be derived by taking the first order partial derivative of Eq. (1) with respect to parameters \mathbf{p}

$$\dot{\mathbf{S}} = \left(\frac{\partial \mathbf{f}}{\partial \mathbf{x}} \right) \mathbf{S} + \left(\frac{\partial \mathbf{f}}{\partial \mathbf{p}} \right) \quad (3)$$

where $\mathbf{S} = \frac{\partial \mathbf{x}}{\partial \mathbf{p}}$ is the state sensitivity matrix. By solving this equation (see Additional file 1 for details) the state sensitivity matrix can be written as

$$\mathbf{S} = \int_{t_0}^t \left(\mathbf{V}(\tau) \left(\frac{\partial \mathbf{f}}{\partial \mathbf{p}} \right) \right) d\tau \quad (4)$$

where $\mathbf{V}(\tau)$ is a matrix computed at time τ . It means that \mathbf{S} has a linear integral relation with the matrix $\left(\frac{\partial \mathbf{f}}{\partial \mathbf{p}} \right)$ from t_0 to t . If at any time $\left(\frac{\partial \mathbf{f}}{\partial \mathbf{p}} \right)$ has the same linearly dependent columns, the corresponding columns in \mathbf{S} will also be linearly dependent, i.e. the corresponding parameters are correlated. Therefore, we can identify parameter correlations by checking the linear dependences of the column in the matrix $\left(\frac{\partial \mathbf{f}}{\partial \mathbf{p}} \right)$ which is composed of the first order partial derivatives of the right-hand side

of the ODEs. Based on Eq. (2), the output sensitivity matrices are, respectively, given by

$$\frac{\partial \mathbf{y}}{\partial \mathbf{p}} = \frac{\partial \mathbf{h}}{\partial \mathbf{x}} \frac{\partial \mathbf{x}}{\partial \mathbf{p}} = - \frac{\partial \mathbf{h}}{\partial \mathbf{x}} \left(\frac{\partial \mathbf{f}}{\partial \mathbf{x}} \right)^{-1} \frac{\partial \mathbf{f}}{\partial \mathbf{p}} \quad (5)$$

$$\frac{\partial \mathbf{y}}{\partial \mathbf{q}} = \frac{\partial \mathbf{h}}{\partial \mathbf{q}} \quad (6)$$

To ensure unique estimation of the parameters (i.e. all parameters to be identifiable), based on the measured data of \mathbf{y} , it is necessary that the columns in the output sensitivity matrices $\frac{\partial \mathbf{y}}{\partial \mathbf{p}}$, $\frac{\partial \mathbf{y}}{\partial \mathbf{q}}$ are linearly independent. From Eq. (6), relations of the columns in $\frac{\partial \mathbf{y}}{\partial \mathbf{q}}$ can be easily detected. The difficulty comes from Eq. (5), since the sensitivity functions in $\frac{\partial \mathbf{y}}{\partial \mathbf{p}}$ cannot be analytically expressed. However, from Eq. (5), the output sensitivity matrix is a linear transformation of $\frac{\partial \mathbf{f}}{\partial \mathbf{p}}$. Consequently, there will be linearly dependent columns in $\frac{\partial \mathbf{y}}{\partial \mathbf{p}}$, if there are linearly dependent columns in $\frac{\partial \mathbf{f}}{\partial \mathbf{p}}$. It means the necessary condition for unique estimation of \mathbf{p} is that, at least, the matrix $\frac{\partial \mathbf{f}}{\partial \mathbf{p}}$ must have a full rank. Based on Eq. (1), $\frac{\partial \mathbf{f}}{\partial \mathbf{p}}$ is expressed as vectors of the first order partial derivative functions

$$\frac{\partial \mathbf{f}}{\partial \mathbf{p}} = \left[\frac{\partial \mathbf{f}}{\partial p_1}, \frac{\partial \mathbf{f}}{\partial p_2}, \dots, \frac{\partial \mathbf{f}}{\partial p_{NP}} \right] \quad (7)$$

Now we analyse relations between the partial derivative functions in Eq. (7). If there is no correlation among the parameters, the columns in Eq. (7) will be linearly independent, i.e. if

$$\alpha_1 \frac{\partial \mathbf{f}}{\partial p_1} + \alpha_2 \frac{\partial \mathbf{f}}{\partial p_2} + \dots + \alpha_{NP} \frac{\partial \mathbf{f}}{\partial p_{NP}} = 0 \quad (8)$$

there must be $\alpha_i = 0, i = 1, \dots, NP$. Otherwise, there will be some groups of vectors in $\frac{\partial \mathbf{f}}{\partial \mathbf{p}}$ which lead to the following cases of linear dependences due to parameter correlations. Let us consider a subset of the parameters with k correlated parameters denoted as $\mathbf{p}_{sub} = [p_{s+1}, p_{s+2}, \dots, p_{s+k}]^T$ with $s + k \leq NP$.

Case 1:

$$\alpha_1 \frac{\partial \mathbf{f}}{\partial p_{s+1}} = \alpha_2 \frac{\partial \mathbf{f}}{\partial p_{s+2}} = \dots = \alpha_k \frac{\partial \mathbf{f}}{\partial p_{s+k}} \quad (9)$$

where $\alpha_i \neq 0, i = 1, \dots, k$. Notice that the coefficient α_i may be a function of the parameters (i.e. $\alpha_i(\mathbf{p})$) and/or of

control inputs (i.e. $\alpha_i(\mathbf{u}(t), \mathbf{p})$). It should be also noted that the control inputs $\mathbf{u}(t)$ are considered as constants in these coefficients, since they will be specified in experimental design. The linear dependences described by Eq. (9) lead to pairwise correlations among the k parameters, i.e. any pair of the parameters in \mathbf{p}_{sub} are correlated. Moreover, the correlations mean a functional relationship between the parameters, i.e. the relationship between the parameters can be expressed by an algebraic equation

$$\phi_{sub}(\gamma(p_{s+1}, p_{s+2}, \dots, p_{s+k})) = 0 \tag{10}$$

where $\gamma(p_{s+1}, p_{s+2}, \dots, p_{s+k})$ denotes a function of the parameters with one set of pairwise correlated parameters. The parameters in this function are compensated each other in an algebraic relationship, e.g. $\gamma(p_{s+1} + p_{s+2} + \dots + p_{s+k})$ or $\gamma(p_{s+1}p_{s+2}\dots p_{s+k})$. Eq. (10) describes the functional relationship between the parameters, e.g. $\phi_{sub}(\gamma(\cdot)) = 1 + \gamma(\cdot) - (\gamma(\cdot))^2 = 0$. Due to the complexity of biological models, an explicit expression of this equation is not available in most cases.

If the coefficients in Eq. (9) are functions of only the parameters, i.e. $\alpha_i(\mathbf{p})$, the parameters are *structurally* non-identifiable. In this case, the correlation relations in Eq. (9) will remain unchanged by specifying *any* values of control inputs. It means that the non-identifiability cannot be remedied through experimental design.

If the coefficients in Eq. (9) are functions of both the parameters and control inputs, i.e. $\alpha_i(\mathbf{u}, \mathbf{p})$, the parameters are *practically* non-identifiable. Different values for \mathbf{u} can be specified which lead to different $\alpha_i(\mathbf{u}, \mathbf{p})$, such that Eq. (9) will not hold and therefore the parameter correlations will be relieved. Since k parameters are correlated, k different values of the control inputs $\mathbf{u}^{(j)}$, ($j = 1, \dots, k$) are required, such that the matrix

$$\frac{\partial \mathbf{f}}{\partial \mathbf{p}_{sub}} = \begin{bmatrix} \frac{\partial \mathbf{f}^{(1)}}{\partial p_{s+1}} & \frac{\partial \mathbf{f}^{(1)}}{\partial p_{s+2}} & \dots & \frac{\partial \mathbf{f}^{(1)}}{\partial p_{s+k}} \\ \frac{\partial \mathbf{f}^{(2)}}{\partial p_{s+1}} & \frac{\partial \mathbf{f}^{(2)}}{\partial p_{s+2}} & \dots & \frac{\partial \mathbf{f}^{(2)}}{\partial p_{s+k}} \\ \vdots & \vdots & \dots & \vdots \\ \frac{\partial \mathbf{f}^{(k)}}{\partial p_{s+1}} & \frac{\partial \mathbf{f}^{(k)}}{\partial p_{s+2}} & \dots & \frac{\partial \mathbf{f}^{(k)}}{\partial p_{s+k}} \end{bmatrix} \tag{11}$$

has a full rank. Notice that the columns in Eq. (11) are only linearly dependent for the same input, but the columns of the whole matrix are linearly independent. In this way, the non-identifiability is remedied. Moreover, a suggestion for experimental design is provided for the

specification of $\mathbf{u}^{(j)}$, ($j = 1, \dots, k$) to obtain k distinct data sets which will be used for parameter estimation.

If all state variables are measurable, according to Eq. (4), this subset of parameters can be uniquely estimated based on the k data sets. If the outputs \mathbf{y} are measured and used for the parameter estimation, it can be concluded from Eq. (5) that at least k data sets are required for unique parameter estimation.

Case 2:

$$\begin{aligned} \alpha_1 \frac{\partial \mathbf{f}}{\partial p_{s+1}} &= \dots = \alpha_{s+l_1} \frac{\partial \mathbf{f}}{\partial p_{s+l_1}}, \quad \dots, \quad \alpha_{s+l_{d-1}+1} \frac{\partial \mathbf{f}}{\partial p_{s+l_{d-1}+1}} \\ &= \dots = \alpha_{s+k} \frac{\partial \mathbf{f}}{\partial p_{s+k}} \end{aligned} \tag{12}$$

and

$$\alpha_{s+k+1} \frac{\partial \mathbf{f}}{\partial p_{s+1}} + \alpha_{s+k+2} \frac{\partial \mathbf{f}}{\partial p_{s+l_1+1}} + \dots + \alpha_{s+k+d} \frac{\partial \mathbf{f}}{\partial p_{s+l_{d-1}+1}} = 0 \tag{13}$$

where $\alpha_i \neq 0, i = 1, \dots, s+k+d$. Similarly, the coefficients may be functions of the parameters and/or of the control inputs. In this case, there are d sets of pairwise correlated parameters (Eq. (12)). A set is not correlated with another set, but all sets are correlated together (Eq. (13)). The functional relationship in this case can be expressed by

$$\phi_{sub}(\gamma_1(p_{s+1}, \dots, p_{s+l_1}), \dots, \gamma_d(p_{s+l_{d-1}+1}, \dots, p_{s+k})) = 0 \tag{14}$$

Based on Eq. (12), the group with the maximum number of parameters $\max(l_1, l_2, \dots, l_d)$ is of importance for data acquisition. From Eq. (13), in the case of practical non-identifiability, data for at least d different inputs is required. The combination of Eqs. (12) and (13) leads to the conclusion that we need a number of $\max(l_1, l_2, \dots, l_d, d)$ data sets with different inputs to eliminate parameter correlations in this case.

Case 3:

$$\alpha_1 \frac{\partial \mathbf{f}}{\partial p_{s+1}} + \alpha_2 \frac{\partial \mathbf{f}}{\partial p_{s+2}} + \alpha_3 \frac{\partial \mathbf{f}}{\partial p_{s+3}} + \dots + \alpha_k \frac{\partial \mathbf{f}}{\partial p_{s+k}} = 0 \tag{15}$$

where $\alpha_i \neq 0, i = 1, \dots, k$. In this case, all k parameters are not pairwise correlated but they are correlated together in one group. The correlation equation in this case is expressed by

$$\phi_s(p_{s+1}, p_{s+1}, \dots, p_{s+k}) = 0 \tag{16}$$

which means there is no set of correlated parameters in this case. The approach described above is able to identify pairwise and higher order parameter correlations in the state equations (Eq. (1)). In the same way, correlations among parameters in \mathbf{q} in the output equations (Eq. (2)) can also be detected based on the first order partial derivative functions in Eq. (6).

From the results of this correlation analysis, the maximum number of correlated parameters of the correlation groups can be detected. This corresponds to the minimum number of data sets required for unique estimation of all parameters in the model. Furthermore, it is noted that the initial state of the model has no impact on the parameter correlations, which means that any initial state can be used for the experimental runs for the data acquisition.

For complex models, the correlation equations (Eqs. (10), (14), (16)) cannot be analytically expressed. A numerical method has to be used to illustrate the relationships of correlated parameters of a given model, which is discussed in the next section.

Interpretation of parameter correlations

Here we give an interpretation of parameter correlations in a biological model. Geometrically, for NP parameters, i.e. $\mathbf{p} = [p_1, p_2, \dots, p_{NP}]^T$, the estimation task can be considered as searching for true parameter values \mathbf{p}^* in the NP -dimensional parameter space. To do this, we need NP linearly independent surfaces in the parameter space which should pass through \mathbf{p}^* . Mathematically, such surfaces are described by linearly independent equations with the unknown parameters. We define such equations based on the results of fitting model Equations (1) to a data set (j) by minimizing the following cost function

$$\min_{\mathbf{p}} F^{(j)}(\mathbf{p}) = \sum_{l=1}^M \sum_{i=1}^n w_{i,l} \left(x_{i,l}^{(j)}(\mathbf{p}) - \hat{x}_{i,l}^{(j)} \right)^2 \quad (17)$$

where M is the number of sampling points, n is the number of state variables and $\hat{\mathbf{x}}$ denotes the measured data while $\mathbf{x}(\mathbf{p})$ the state variables predicted by the model. $w_{i,l}$ are weighting factors. The fitting results will depend on the data set resulted from the control inputs $\mathbf{u}^{(j)}$, the values of $w_{i,l}$, and the noise level of the measured data. For a geometric interpretation of parameter correlations, we assume to use idealized measurement data, i.e. data without any noises. Based on this assumption, the residual function (17) should be zero, when the true parameter set \mathbf{p}^* is applied, i.e.

$$x_{i,l}^{(j)}(\mathbf{p}^*) - \hat{x}_{i,l}^{(j)} = 0, \quad i = 1, \dots, n, \quad l = 1, \dots, M \quad (18)$$

It is noted that Eq. (18) is true for any noise-free data set employed for the fitting and independent of $w_{i,q}$. Now we define a zero residual equation (ZRE) as

$$\phi_{i,l}^{(j)}(\mathbf{p}) = x_{i,l}^{(j)}(\mathbf{p}) - \hat{x}_{i,l}^{(j)} = 0 \quad (19)$$

This equation contains the parameters as unknowns and corresponds to a zero residual surface passing through the true parameter point \mathbf{p}^* . It means that a zero residual surface is built by parameter values which lead to a zero residual value. This suggests that we can find \mathbf{p}^* by solving NP linearly independent ZREs. From the first order Taylor expansion of Eq. (19), the linear dependences of ZREs can be detected by the columns in the following matrix

$$\frac{\partial \mathbf{x}^{(j)}}{\partial \mathbf{p}} = \left[\frac{\partial \mathbf{x}^{(j)}}{\partial p_1}, \frac{\partial \mathbf{x}^{(j)}}{\partial p_2}, \dots, \frac{\partial \mathbf{x}^{(j)}}{\partial p_{NP}} \right] \quad (20)$$

where $\mathbf{x}^{(j)} = [x_{1,1}^{(j)}, x_{1,2}^{(j)}, \dots, x_{1,M}^{(j)}, \dots, x_{n,1}^{(j)}, x_{n,2}^{(j)}, \dots, x_{n,M}^{(j)}]^T$. Eq. (20) is exactly the state sensitivity matrix calculated by fitting to the given data set (j). This means, under the idealized data assumption, a zero residual value delivered after the fitting is associated to surfaces passing through the true parameter point. When there are no parameter correlations, the number of linearly independent ZREs will be greater than NP and thus the true parameter point can be found by fitting the current data set.

If there are parameter correlations, the fitting will lead to zero residual surfaces in the subspace of the correlated parameters. For instance, for a group of k correlated parameters, the zero residual surfaces (Eq. (19)) will be reduced to a single ZRE represented by Eq. (10), Eq. (14), or Eq. (16). Therefore, in the case of *practical* non-identifiability, k data sets are needed to generate k linearly independent ZREs so that the k parameters can be uniquely estimated. In the case of *structural* non-identifiability, the correlated relations are independent of data sets. It means fitting different data sets will lead to the same ZRE and thus the same surfaces in the parameter subspace.

If the measured data are with noises, the fitting results will lead to a nonzero residual value and nonzero residual surfaces, i.e.

$$\phi_{i,l}^{(j)}(\mathbf{p}) = x_{i,l}^{(j)}(\mathbf{p}) - \hat{x}_{i,l}^{(j)} = \varepsilon_{i,l} \quad (21)$$

where $\varepsilon_{i,l} \neq 0$. Thus the nonzero residual surfaces will not pass through the true parameter point. However, based on Eq. (20) and Eq. (21) the first order partial derivatives remain unchanged. It means that parameter correlations do not depend on the quality of the measured data. Moreover, it can be seen from Eq. (19) and Eq. (21) that the zero residual surfaces and the nonzero residual surfaces will be parallel in the subspace of the correlated parameters.

Results and discussion

We consider a three-step pathway modelled by 8 nonlinear ordinary differential equations (ODEs) containing 8 metabolic concentrations (state variables) and 36 parameters [30-32], as given in Eqs. (22-29). The P and S values in the ODEs are considered as two control inputs specified by experimental design. No output equations were considered for this model in the previous studies.

$$\dot{x}_1 = \frac{p_1}{1 + \left(\frac{P}{p_2}\right)^{p_3} + \left(\frac{P_4}{S}\right)^{p_5}} - p_6 x_1 \quad (22)$$

$$\dot{x}_2 = \frac{p_7}{1 + \left(\frac{P}{p_8}\right)^{p_9} + \left(\frac{p_{10}}{x_7}\right)^{p_{11}}} - p_{12} x_2 \quad (23)$$

$$\dot{x}_3 = \frac{p_{13}}{1 + \left(\frac{P}{p_{14}}\right)^{p_{15}} + \left(\frac{p_{16}}{x_8}\right)^{p_{17}}} - p_{18} x_3 \quad (24)$$

$$\dot{x}_4 = \frac{p_{19} x_1}{p_{20} + x_1} - p_{21} x_4 \quad (25)$$

$$\dot{x}_5 = \frac{p_{22} x_2}{p_{23} + x_2} - p_{24} x_5 \quad (26)$$

$$\dot{x}_6 = \frac{p_{25} x_3}{p_{26} + x_3} - p_{27} x_6 \quad (27)$$

$$\dot{x}_7 = \frac{p_{28} x_4 (S - x_7)}{p_{29} \left(1 + \frac{S}{p_{29}} + \frac{x_7}{p_{30}}\right)} - \frac{p_{31} x_5 (x_7 - x_8)}{p_{32} \left(1 + \frac{x_7}{p_{32}} + \frac{x_8}{p_{33}}\right)} \quad (28)$$

$$\dot{x}_8 = \frac{p_{31} x_5 (x_7 - x_8)}{p_{32} \left(1 + \frac{x_7}{p_{32}} + \frac{x_8}{p_{33}}\right)} - \frac{p_{34} x_6 (x_8 - P)}{p_{35} \left(1 + \frac{x_8}{p_{35}} + \frac{P}{p_{36}}\right)} \quad (29)$$

This pathway model was studied by Moles et al. [31] using 16 noise-free data sets and Rodriguez-Fernandez et al. [32] using 16 both noise-free and noisy data sets, respectively. They showed some strong parameter correlations in several groups. Accurate parameter values were identified in [32]. However, a clear correlation analysis of the parameters and the relationship between the parameter correlations and the numbers of data sets with different inputs required for the parameter estimation were not given in the previous studies.

Identification of correlations

Now we identify parameter correlations in this model using our approach. Given the model represented by Eqs. (22-29), the first order partial derivative functions

can be readily derived leading to the following linear dependences (see Additional file 1 for detailed derivation).

From Eq. (22),

$$\alpha_1 \frac{\partial f_1}{\partial p_1} = \alpha_2 \frac{\partial f_1}{\partial p_2} = \dots = \alpha_5 \frac{\partial f_1}{\partial p_5} \quad (30)$$

From Eq. (23),

$$\alpha_6 \frac{\partial f_2}{\partial p_8} = \frac{\partial f_2}{\partial p_9} \quad \text{and} \quad \alpha_7 \frac{\partial f_2}{\partial p_7} + \alpha_8 \frac{\partial f_2}{\partial p_{10}} = \frac{\partial f_2}{\partial p_8} \quad (31)$$

From Eq. (24),

$$\alpha_9 \frac{\partial f_3}{\partial p_{14}} = \frac{\partial f_3}{\partial p_{15}} \quad \text{and} \quad \alpha_{10} \frac{\partial f_3}{\partial p_{13}} + \alpha_{11} \frac{\partial f_3}{\partial p_{16}} = \frac{\partial f_3}{\partial p_{14}} \quad (32)$$

From Eq. (28),

$$\alpha_{12} \frac{\partial f_7}{\partial p_{28}} + \alpha_{13} \frac{\partial f_7}{\partial p_{29}} = \frac{\partial f_7}{\partial p_{30}} \quad (33)$$

From Eq. (29),

$$\alpha_{14} \frac{\partial f_8}{\partial p_{35}} = \frac{\partial f_8}{\partial p_{36}} \quad (34)$$

The coefficients in Eqs. (30) – (34), α_i ($i = 1, \dots, 14$), are functions of the corresponding parameters and controls in the individual state equations (see Additional file 1). Based on these results, correlated parameters in this model can be described in 5 groups:

Group 1: $G_1(p_1, p_2, p_3, p_4, p_5)$, among which any pair of parameters are pairwise correlated;

Group 2: $G_2(p_7, p_8, p_9, p_{10})$, among which p_8, p_9 are pairwise correlated and p_7, p_8, p_{10} as well as p_7, p_9, p_{10} are correlated, respectively.

Group 3: $G_3(p_{13}, p_{14}, p_{15}, p_{16})$, among which p_{14}, p_{15} are pairwise correlated and p_{13}, p_{14}, p_{16} as well as p_{13}, p_{15}, p_{16} are correlated, respectively.

Group 4: $G_4(p_{28}, p_{29}, p_{30})$, the parameters are correlated together but not pairwise;

Group 5: $G_5(p_{35}, p_{36})$, they are pairwise correlated.

Since the coefficients are functions of both of the parameters and the control inputs, these correlated parameters are practically non-identifiable for a single set of data. It is noted that, in G_2 and G_3 , the maximum number of correlated parameters is three. Among the 5 correlated parameter groups the maximum number of correlated parameters is 5 (from G_1). It means at least 5

data sets with different control values are required to uniquely estimate the 36 parameters of this model.

Verification of the correlations by fitting the model

To verify the proposed approach and check the correlations in this model, we carried out numerical

experiments by fitting the parameters to a certain number of simulated data sets with different inputs. The fitting method used is a modified sequential approach suitable for handling multiple data sets [33,34].

We used the nominal parameter values given in [31], initial state values as well as *P* and *S* values (see Additional file 1)

Table 1 Fitted parameter values based on different data sets

No.	P^*	$P^{(1)}$	$P^{(1)+(2)}$	$P^{(1)+(2)+(3)}$	$P^{(1)+\dots+(4)}$	$P^{(1)+\dots+(5)}$	$P^{(1)+\dots+(5)}(w)$
1(G_1)	1.0	1.06763	1.07763	1.60486	1.73180	1.00000	0.97145
2(G_1)	1.0	1.40146	0.91495	0.82116	0.75989	0.99998	1.05917
3(G_1)	2.0	1.47116	1.16323	2.39189	2.00001	2.00006	1.86755
4(G_1)	1.0	1.55173	1.01042	2.30123	3.19504	1.00000	0.98664
5(G_1)	2.0	1.40069	1.24912	0.32136	0.25317	2.00000	2.01339
6	1.0	1.00000	1.00002	1.00000	1.00000	1.00000	0.98154
7(G_2)	1.0	1.00927	1.02815	1.00000	1.00000	1.00000	0.99124
8(G_2)	1.0	1.32173	0.95504	1.00000	1.00000	1.00000	0.99919
9(G_2)	2.0	1.34185	1.18286	2.00000	2.00000	2.00000	1.93527
10(G_2)	1.0	1.00477	1.01393	1.00000	1.00000	1.00000	0.98693
11	2.0	1.99973	2.00007	2.00000	2.00000	2.00000	2.03582
12	1.0	0.99944	1.00019	1.00000	1.00000	1.00000	1.00435
13(G_3)	1.0	1.00572	1.05126	1.00001	1.00001	1.00001	1.03448
14(G_3)	1.0	1.39147	0.90768	1.00000	1.00000	1.00000	0.99558
15(G_3)	2.0	1.45117	1.00760	2.00003	2.00002	2.00001	1.98699
16(G_3)	1.0	1.00280	1.02531	1.00001	1.00000	1.00001	0.99786
17	2.0	1.99987	1.99999	1.99999	1.99999	1.99999	1.99586
18	1.0	1.00016	1.00000	1.00000	1.00000	1.00000	1.03924
19	0.1	0.10016	0.10000	0.10000	0.10000	0.10000	0.10000
20	1.0	1.00263	1.00000	1.00000	1.00000	1.00001	0.99469
21	0.1	0.10003	0.10000	0.10000	0.10000	0.10000	0.10007
22	0.1	0.10010	0.10000	0.10000	0.10000	0.10000	0.10000
23	1.0	1.00127	1.00000	1.00000	1.00000	1.00000	0.99581
24	0.1	0.10003	0.10000	0.10000	0.10000	0.10000	0.10025
25	0.1	0.10003	0.10000	0.10000	0.10000	0.10000	0.10492
26	1.0	1.00023	1.00002	1.00001	1.00000	1.00001	1.05077
27	0.1	0.10001	0.10000	0.10000	0.10000	0.10000	0.10120
28(G_4)	1.0	0.96519	0.99594	1.00000	1.00000	1.00000	1.01865
29(G_4)	1.0	1.62390	1.04672	1.00000	1.00000	1.00001	0.90507
30(G_4)	1.0	1.56817	1.04245	1.00000	0.99999	1.00000	0.85521
31	1.0	0.99997	1.00000	1.00000	1.00000	1.00000	1.11984
32	1.0	1.00110	1.00000	1.00000	1.00000	1.00000	0.97161
33	1.0	1.00207	0.99998	1.00000	0.99998	0.99998	1.33808
34	1.0	0.99956	1.00000	1.00000	1.00000	1.00000	1.01811
35(G_5)	1.0	1.05000	1.00001	1.00000	1.00000	1.00000	1.05077
36(G_5)	1.0	2.03075	0.99999	1.00000	1.00000	1.00000	1.20947
Residual value		3.62E-9	4.26E-9	5.31E-9	6.49E-9	5.35E-9	1.12E-0

P^* are the nominal (true) values, $P^{(1)}$ the values based on the 1st data set, $P^{(1)+(2)}$ based on the 1st, 2nd data sets together, $P^{(1)+(2)+(3)}$ based on the 1st, 2nd, and 3rd data sets, $P^{(1)+\dots+(4)}$ based on the 1st to 4th data sets, and $P^{(1)+\dots+(5)}$ based on the 5 data sets, respectively. (w) means results from 10% noises on the data. Correlated parameter groups are highlighted separately.

given in [32] to generate 5 noise-free data sets with different inputs containing the time courses of the 8 state variables. For each data set 120 data points were taken with 1 minute as sampling time.

For fitting the parameters we used random values for all 36 parameters to initialize the computation and all weights in Eqn. (17) were set to 1.0. The results were taken by a threshold of the total residual value in the order of 10^{-9} when using noise-free data sets (see Table 1).

Figure 1A (upper panel) shows the angles between the columns of the state sensitivity matrix by fitting to the 1st data set. The zero angles (red lines) mean that the corresponding columns are pairwise parallel. According to Figure 1A, 4 pairwise correlated parameter groups (i.e. $(p_1, p_2, p_3, p_4, p_5)$, (p_8, p_9) , (p_{14}, p_{15}) , (p_{35}, p_{36})) can be detected. However, these are not the same results as identified by the analysis of the model equations. This is because a dendrogram only shows pairwise correlations; it cannot detect higher order interrelationships among the parameters.

To illustrate the geometric interpretation, we first take the group of $G_5(p_{35}, p_{36})$ as an example to construct ZREs, i.e. to plot the correlated relations between p_{35} and p_{36} . This was done by repeatedly fitting the model to the 5 individual data sets with different inputs, respectively, with fixed values of p_{35} . The resulting 5 zero residual surfaces (lines) in the subspace of p_{35} and p_{36} are shown in Figure 2A. As expected, the zero residual surfaces resulted from different data sets cross indeed at the true parameter point in the parameter subspace. Figure 2B shows the relations between p_{35} and p_{36} by fitting the parameters separately to the same 5

data sets on which a Gaussian distributed error of 10% was added. It can be seen that, due to the measurement noises, the crossing points of the nonzero residual surfaces are at different positions but near the true parameter point. Moreover, by comparing the lines in Figure 2A with Figure 2B, it can be seen that the corresponding zero residual surfaces and nonzero residual surfaces are indeed parallel, when fitting the same data set without noises or with noises, respectively.

Figure 3 shows the residual surfaces based on fitting to 2 individual noise-free data sets (Figure 3A) and to the same 2 data sets together (Figure 3B). It is shown from Figure 3A that, due to the correlation, two hyperbolic cylinders are built by separately fitting to individual data sets. The bottom minimum lines of the two cylinders corresponding to the zero residual value cross at the true parameter point. Fitting to the two data sets together leads to an elliptic paraboloid (Figure 3B) which has only one minimum point with the zero residual value. This point is the true parameter point, which means the remedy of the correlation between p_{35} and p_{36} .

Since the maximum number of parameters among the correlation groups is 5, according to our approach, at least 5 data sets with different inputs are needed to uniquely determine the parameter set. The last column in Table 1 ($\mathbf{P}^{(1)+\dots+(5)}$) shows the parameter values from fitting the model to the 5 data sets together. It can be seen that all of the 36 parameter values fitted are almost at their true values. According our geometric interpretation, this means that the 5 zero residual surfaces expanded by together fitting to the 5 data sets cross at the true parameter point in the parameter subspace.

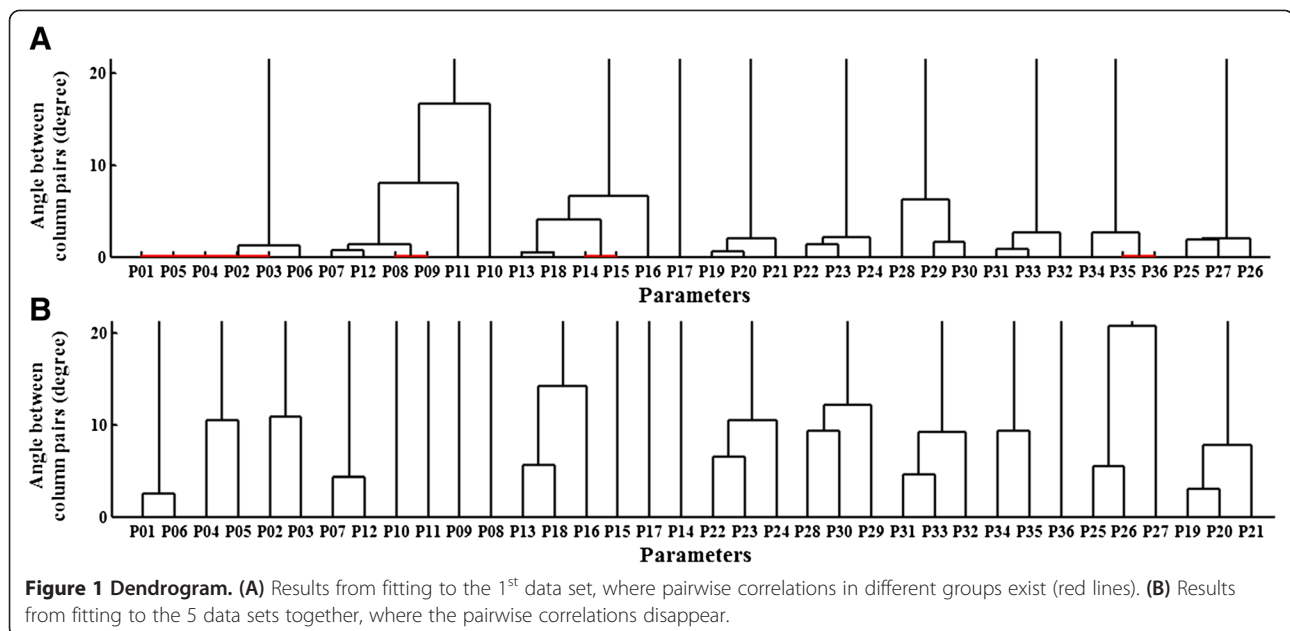


Figure 1 Dendrogram. (A) Results from fitting to the 1st data set, where pairwise correlations in different groups exist (red lines). (B) Results from fitting to the 5 data sets together, where the pairwise correlations disappear.

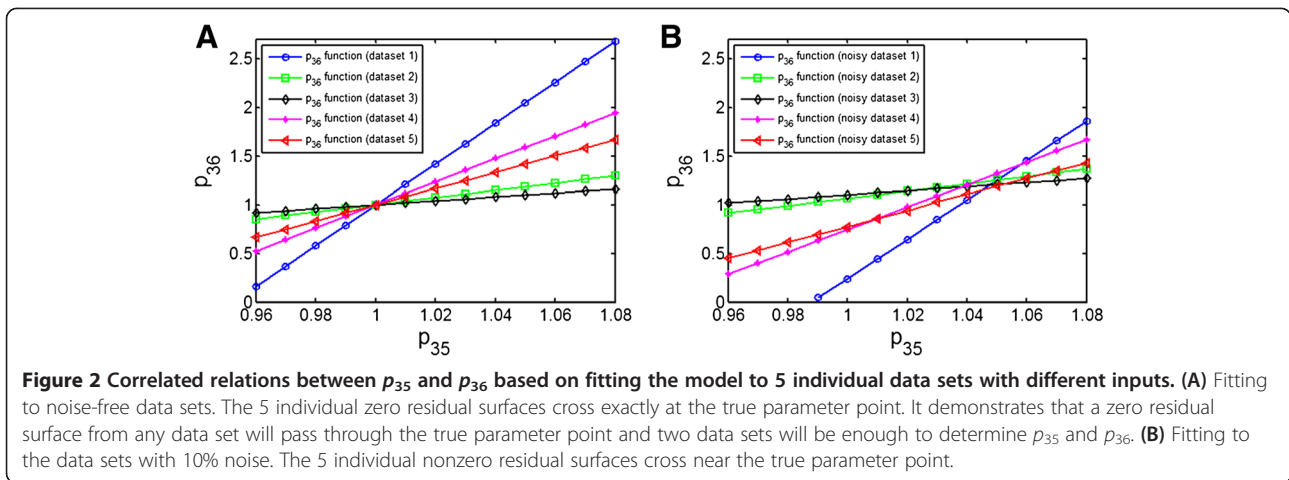
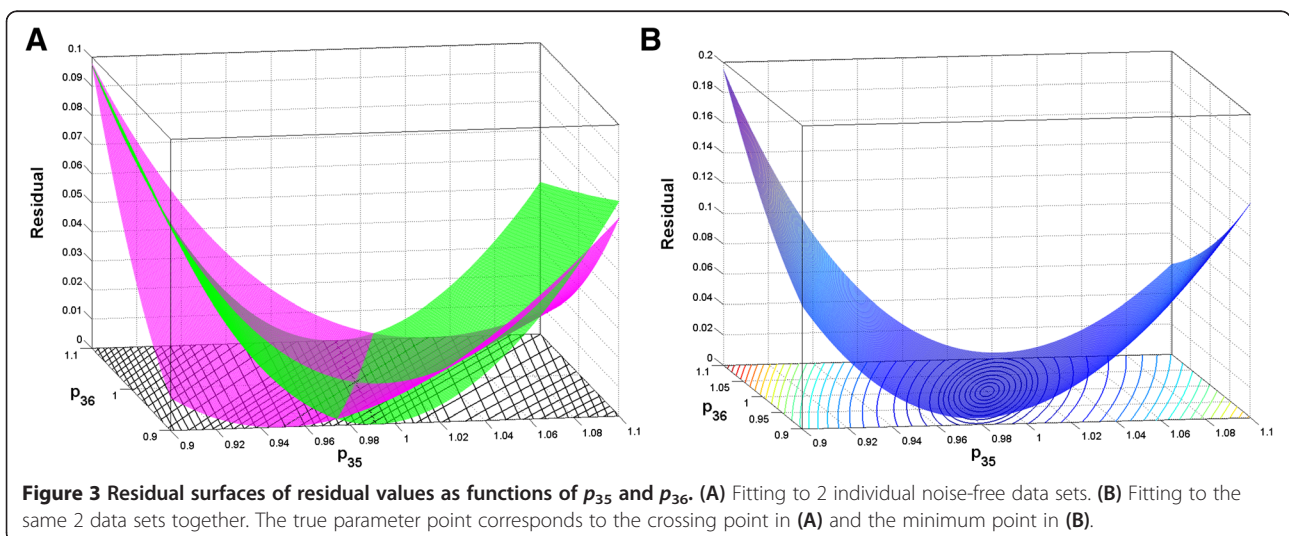


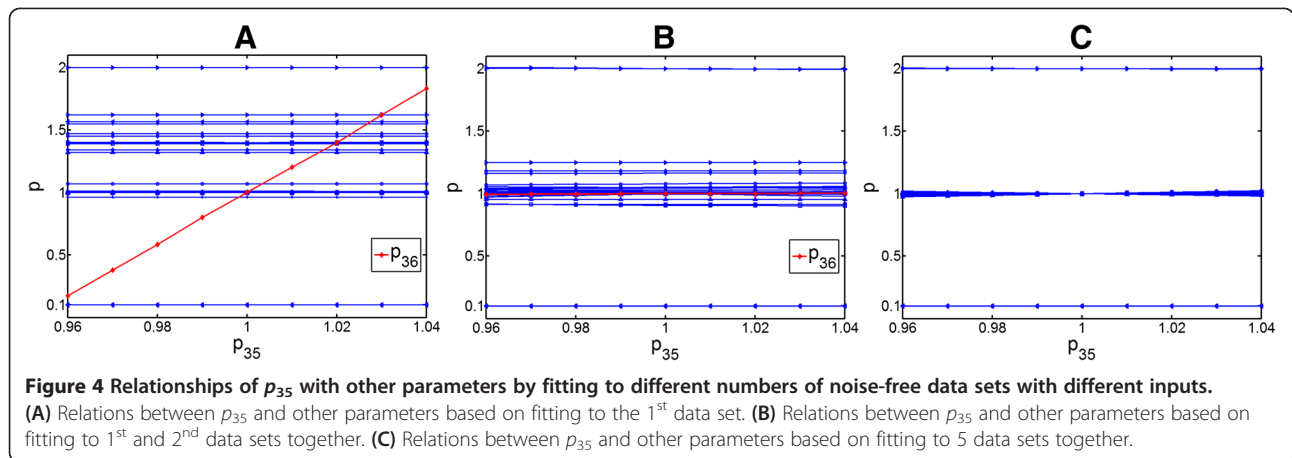
Figure 1B (lower panel) shows these correlated relations indeed disappear based on the results of fitting to the 5 data sets together.

Moreover, it is shown in Table 1 ($\mathbf{P}^{(1)+(2)}$) that the correlation between p_{35} and p_{36} can be remedied by fitting two data sets together. As expected, it can be seen that in $\mathbf{P}^{(1)+(2)}$ the parameters in G_1 are not well fitted (i.e. 5 correlated parameters cannot be uniquely determined by two data sets). It is also interesting to see in $\mathbf{P}^{(1)+(2)}$ the parameter values in G_2 , G_3 and G_4 are also not well estimated. This is because the degree of freedom of $G_2(p_7, p_8, p_9, p_{10})$, $G_3(p_{13}, p_{14}, p_{15}, p_{16})$, and $G_4(p_{28}, p_{29}, p_{30})$ is 3. Indeed, as shown in Table 1 ($\mathbf{P}^{(1)+(2)+(3)}$), these parameters are exactly determined based on fitting the model to 3 data sets together. However, it is shown in Table 1 from the parameter values of $\mathbf{P}^{(1)+(2)+(3)}$ and $\mathbf{P}^{(1)+\dots+(4)}$ that a number of data sets less than 5 is not enough to remedy the correlations of the parameters in G_1 .

To test the sensitivity of the parameter results to measurement errors, we also fitted the model to the same 5 data sets with different inputs and with 10% noise level together. As shown in the last column in Table 1 ($\mathbf{P}^{(1)+\dots+(5)(w)}$), to some extent, the parameter values identified are deviated from the true values due to an increased residual value. But the overall parameter quality is quite good. It means the crossing points of the 5 nonzero residual surfaces expanded by the 5 noisy data sets are quite close to the true parameter point.

Figure 4 shows profiles of all parameters as a function of p_{35} , based on different number of data sets used for fitting. It is seen from Figure 4A that only p_{36} is correlated with p_{35} (red line). Moreover, it can be seen that, by fitting to one data set, the other parameters which have higher order interrelationships in other groups cannot be well determined. As shown in Figure 4B, the correlation between p_{35} and p_{36} is remedied by fitting to





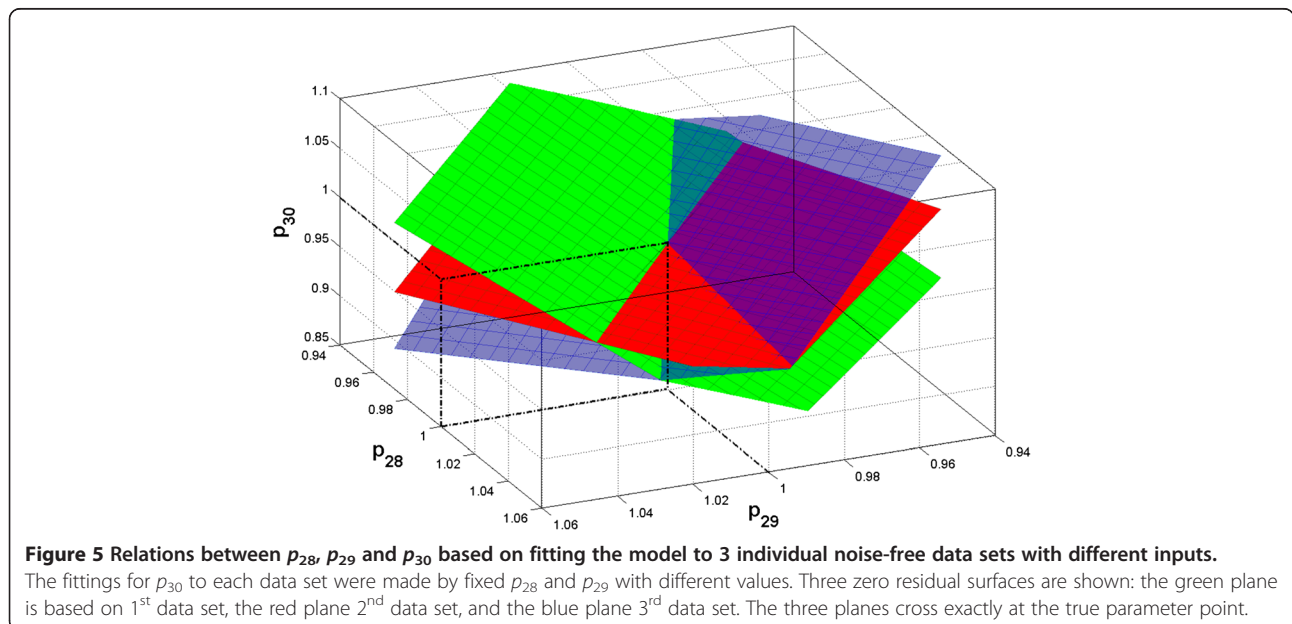
two data sets together and, moreover, the parameters tend to approach their true values (i.e. 0.1, 1.0 and 2.0, see Table 1). Finally, all parameters are uniquely determined (i.e. clearly at the three true values), when 5 data sets were used together for fitting the model, as shown in Figure 4C.

These results clearly demonstrate the scope of our approach to identifying parameter correlations. Moreover, it is clearly seen that adding more data sets with different inputs can remedy the parameter non-identifiability problem in some complex models, but a necessary number of data sets with different inputs (5 for this example) is enough.

To illustrate a higher order interrelationship among parameters, estimations were made by separately fitting the model to 3 individual data sets to plot the

relations of the parameters in $G_4(p_{28}, p_{29}, p_{30})$, as shown in Figure 5. It can be seen that the three zero residual surfaces (planes) resulted from the three individual data sets cross exactly at the true parameter point in the subspace of the 3 parameters. This demonstrates our geometric interpretation of parameter correlations, i.e. to estimate a group of three correlated parameters at least three distinct data sets with different inputs are needed.

Since parameter correlations determined from the proposed approach are based on the structure of the state equations, our result provides a minimum number of different data sets with different inputs necessary for unique parameter estimation (5 in this example). This is definitely true, if all state variables (8 in this example) are measurable and included in the 5 data sets.



The results shown above are from the solutions of the parameter estimation problem based on the data sets composed of *all* 8 state variables. It is demonstrated that at least 5 data sets with different inputs will be needed to uniquely estimate the 36 parameters. However, our method does not give information on how many state variables which may be fewer than 8 but sufficient to identify the 36 parameters. To achieve this information, we tried to estimate the parameters based on the generated 5 data sets which include fewer measured state variables (as output variables). We checked the identifiability when the 5 data sets consist of data profiles of only a part of the state variables. Computational tests were carried out based on different combinations of the state variables included in the data sets. Table 2 shows the minimum sets of state variables which should be included in the data sets so as to achieve a successful fitting. It can be seen, for instance, the 36 parameters can be uniquely estimated in the case that only the first three state variables (i.e. x_1, x_2, x_3) are included in the 5 data sets. Moreover, the generated data profiles of x_7 and x_8 are also enough for identifying the 36 parameters. Due to insufficient data, estimation runs with fewer numbers of the state variables than listed in Table 2 could not converge, i.e. the parameters will be non-identifiable.

Conclusions

It is well recognized that parameters in many biological models are correlated. Finding the true parameter point remains as a challenge since it is hidden in these correlated relations. In many cases, a direct analysis of parameter correlations based on the output sensitivity matrix depends on experimental design, and the analytical relationship cannot be seen. Instead, we presented a method to analyse parameter correlations based on the matrix of the first order partial derivative functions of state equations which can be analytically derived. In this way, pairwise correlations and higher order interrelationships

among the parameters can be detected. The result gives the information about parameter correlations and thus about the identifiability of parameters when all state variables are measurable for fitting the parameters. Since the output sensitivity matrix is a linear transformation of the matrix of first order partial derivative functions, our correlation analysis approach provides a necessary (but not sufficient) condition of parameter identifiability. That is, if there exist parameter correlations, the corresponding parameters are non-identifiable.

In addition, we introduced residual surfaces in the parameter subspace to interpret parameter correlations. Any point on a zero residual surface will result in a zero residual value. The crossing point of multiple zero residual surfaces leads to the true parameter point. Zero residual surfaces correspond to ZREs resulted from noise-free data sets used for fitting the parameters. If the ZREs are linearly independent (i.e. there are no correlations), the model parameters are identifiable, and otherwise they are non-identifiable. If more linearly independent ZREs can be constructed by adding new data sets with different inputs, the parameters are practically non-identifiable, otherwise they are structurally non-identifiable. In the case of practical non-identifiability the true parameter values can be found by together fitting the model to a necessary number of data sets which is the maximum number of parameters among the correlation groups. If the available measured data are from output variables, this should be regarded as the minimum number of data sets with different inputs required for unique parameter estimation. The results of the case study demonstrate the feasibility of our approach.

Moreover, an interesting result of our approach is that parameter correlations are not affected by the initial state. This means that, experimental runs can be conducted with any initial state to obtain the required data sets with different inputs. More interestingly, according to this result, different data sets with different inputs can be gained in one experimental run by changing the values of the control inputs. It is noted that the proposed approach does not address the identifiability issue of the initial states which would be a future research aspect.

The result of identifiable parameters determined by the proposed approach is theoretical. This means that the quality of the available data (the noise level, the length of sampling time, etc.) has an important impact on the identifiability issue. Parameters which are theoretically identifiable may not be identifiable by an estimator due to low quality of the data. Non-identifiability issues caused by relative data are not considered in this paper. In addition, the identification of parameter correlations based on the output equations is not considered in this paper.

Table 2 Measurable variable sets for a successful fitting

No.	Measured variables
y_1	(x_1, x_2, x_3)
y_2	(x_1, x_2, x_6)
y_3	(x_1, x_3, x_5)
y_4	(x_1, x_5, x_6)
y_5	(x_2, x_4, x_6)
y_6	(x_4, x_5, x_6)
y_7	(x_7, x_8)

Different sets of state variables were used as measurable output variables included in the 5 data sets, respectively. This table shows the groups of a minimum number of state variables used as outputs for the parameter estimation which leads to the convergence to the true parameter point.

Additional file

Additional file 1: Derivation of the sensitivity matrix, partial derivative functions of the case study, and the values of control inputs for generating data sets in the case study.

Competing interests

The authors declare that they have no competing interests.

Author's contributions

PL developed the methodology, wrote the manuscript and supervised the study. QDV wrote the software and designed the study. Both authors read and approved the final manuscript.

Acknowledgements

We thank M. Bartl, C. Kaleta, R. Guthke and H. M. La for carefully reading the manuscript and their suggestions for improving the description. This work has been supported by the PhD scholarship from Vietnamese government to QDV. In addition, we thank the anonymous reviewers for their comments that improved the manuscript.

Received: 14 May 2013 Accepted: 12 September 2013

Published: 22 September 2013

References

1. Ashyraliyev M, Fomekong-Nanfack Y, Kaandorp JA, Blom JG: **Systems biology: parameter estimation for biochemical models.** *FEBS J* 2009, **276**:886–902.
2. Chou IC, Voit EO: **Recent developments in parameter estimation and structure identification of biochemical and genomic systems.** *Math Biosci* 2009, **219**:57–83.
3. Jacquéz JA: **Design of experiments.** *J Franklin Inst* 1998, **335B**:259–279.
4. Kreutz C, Timmer J: **Systems biology: experimental design.** *FEBS J* 2009, **276**:923–942.
5. Bauer I, Bock HG, Körkel S, Schlöder JP: **Numerical methods for optimum experimental design in DAE systems.** *J Comp Appl Math* 2000, **120**:1–25.
6. Peifer M, Timmer J: **Parameter estimation in ordinary differential equations for biological processes using the method of multiple shooting.** *IET Syst Biol* 2007, **1**:78–88.
7. Zavala VM, Laird CD, Biegler LT: **Interior-point decomposition approaches for parallel solution of large-scale nonlinear parameter estimation problems.** *Chem Eng Sci* 2008, **63**:4834–4845.
8. Esposito WR, Floudas CA: **Global optimization for the parameter estimation of differential-algebraic systems.** *Ind Eng Chem Res* 2000, **39**:1291–1310.
9. Miró A, Pozo C, Guillén-Gosálbez G, Egea JA, Jiménez L: **Deterministic global optimization algorithm based on outer approximation for the parameter estimation of nonlinear dynamic biological systems.** *BMC Bioinformatics* 2012, **13**:90.
10. Gonzalez OR, Küper C, Jung K, Naval PC Jr, Mendoza E: **Parameter estimation using simulated annealing for S-system models of biological networks.** *Bioinformatics* 2007, **23**:480–486.
11. Kikuchi S, Tominaga D, Arita M, Takahashi K, Tomita M: **Dynamic modelling of genetic networks using genetic algorithm and S-system.** *Bioinformatics* 2003, **19**:643–650.
12. Rodríguez-Fernandez M, Egea JA, Banga JR: **Novel metaheuristic for parameter estimation in nonlinear dynamic biological systems.** *BMC Bioinformatics* 2006, **7**:483.
13. Gutenkunst RN, Waterfall JJ, Casey FP, Brown KS, Myers CR, Sethna JP: **Universally sloppy parameter sensitivities in systems biology models.** *PLoS Comp Biol* 2007, **3**:1871–1878.
14. Raue A, Kreutz C, Maiwald T, Klingmüller U, Timmer J: **Addressing parameter identifiability by model-based experimentation.** *IET Syst Biol* 2011, **5**:120–130.
15. Raue A, Kreutz C, Maiwald T, Bachmann J, Schilling M, Klingmüller U, Timmer J: **Structural and practical identifiability analysis of partially observable dynamical models by exploiting the profile likelihood.** *Bioinformatics* 2009, **25**:1923–1929.
16. McLean KAP, McAuley KB: **Mathematical modelling of chemical processes-obtaining the best model predictions and parameter estimates using identifiability and estimability procedures.** *Can J Chem Eng* 2012, **90**:351–365.
17. Chappel MJ, Godfrey KR, Vajda S: **Global identifiability of the parameters of nonlinear systems with specified inputs: a comparison of methods.** *Math Biosci* 1990, **102**:41–73.
18. Meshkat N, Eisenberg M, DiStefano JJ III: **An algorithm for finding globally identifiable parameter combinations of nonlinear ODE models using Gröbner bases.** *Math Biosci* 2009, **222**:61–72.
19. Chis OT, Banga JR, Balsa-Canto E: **Structural identifiability of systems biology models: a critical comparison of methods.** *PLoS ONE* 2011, **6**:e27755.
20. Vajda S, Rabitz H, Walter E, Lecourtier Y: **Qualitative and quantitative identifiability analysis of nonlinear chemical kinetic models.** *Chem Eng Comm* 1989, **83**:191–219.
21. Yao KZ, Shaw BM, Kou B, McAuley KB, Bacon DW: **Modeling ethylene/butene copolymerization with multi-site catalysts: parameter estimability and experimental design.** *Polym React Eng* 2003, **11**:563–588.
22. Chu Y, Jayaraman A, Hahn J: **Parameter sensitivity analysis of IL-6 signalling pathways.** *IET Syst Biol* 2007, **1**:342–352.
23. Cintrón-Arias A, Banks HT, Capaldi A, Lloyd AL: **A sensitivity matrix based methodology for inverse problem formulation.** *J Inv Ill-Posed Problems* 2009, **17**:545–564.
24. Dobre S, Bastogne T, Profeta C, Barberi-Heyob M, Richard A: **Limits of variance-based sensitivity analysis for non-identifiability testing in high dimensional dynamic models.** *Automatica* 2012, **48**:2740–2749.
25. Steiert B, Raue A, Timmer J, Kreutz C: **Experimental design for parameter estimation of gene regulatory networks.** *PLoS ONE* 2012, **7**:e40052.
26. Hengl S, Kreutz C, Timmer J, Maiwald T: **Data-based identifiability analysis of nonlinear dynamical models.** *Bioinformatics* 2007, **23**:2612–2618.
27. Bachmann J, Raue A, Schilling M, Böhm ME, Kreutz C, Kaschek D, Busch H, Gretz N, Lehmann WD, Timmer J, Klingmüller U: **Division of labor by dual feedback regulators controls JAK2/STAT5 signaling over broad ligand range.** *Mol Sys Bio* 2011, **7**:516.
28. Chu Y, Hahn J: **Parameter set selection for estimation of nonlinear dynamic systems.** *AIChE J* 2007, **53**(11):2858–2870.
29. Quaiser T, Mönningmann M: **Systematic identifiability testing for nambiguous mechanistic modeling – application to JAK-STAT, MAP kinase, and NF-κB signaling pathway models.** *BMC Syst Biol* 2009, **3**:50.
30. Mendes P: **Modeling large biological systems from functional genomic data: parameter estimation.** In *Foundations of Systems Biology.* Cambridge MA: MIT Press; 2001:163–186.
31. Moles CG, Mendes P, Banga JR: **Parameter estimation in biochemical pathways: a comparison of global optimization methods.** *Genome Res* 2003, **13**:2467–2474.
32. Rodríguez-Fernandez M, Mendes P, Banga JR: **A hybrid approach for efficient and robust estimation in biochemical pathways.** *Biosystems* 2006, **83**:248–265.
33. Faber R, Li P, Wozny G: **Sequential parameter estimation for large-scale systems with multiple data sets. I: computational framework.** *Ind Eng Chem Res* 2003, **42**:5850–5860.
34. Zhao C, Vu QD, Li P: **A quasi-sequential parameter estimation for nonlinear dynamic systems based on multiple data profiles.** *Korean J Chem Eng* 2013, **30**:269–277.

doi:10.1186/1752-0509-7-91

Cite this article as: Li and Vu: Identification of parameter correlations for parameter estimation in dynamic biological models. *BMC Systems Biology* 2013 7:91.

Additional file 1 for

Identification of parameter correlations for parameter estimation in dynamic biological models

Derivation of the sensitivity matrix, partial derivative functions of the case study, and the values of control inputs for generating data sets in the case study.

1) The sensitivity matrix derivation

Consider the sensitivity equation

$$\dot{S} = \left(\frac{\partial f}{\partial x} \right) S + \left(\frac{\partial f}{\partial p} \right) \quad (\text{A1})$$

Using the explicit Euler method at time point $t = k \Delta t$ with a small time interval Δt , we can write Eq. (A1) in the discrete form

$$\frac{S_k - S_{k-1}}{\Delta t} = \left(\frac{\partial f}{\partial x} \right)_{k-1} S_{k-1} + \left(\frac{\partial f}{\partial p} \right)_{k-1} \quad (\text{A2})$$

It leads to

$$S_k = S_{k-1} + \Delta t \left(\frac{\partial f}{\partial x} \right)_{k-1} S_{k-1} + \Delta t \left(\frac{\partial f}{\partial p} \right)_{k-1} = \left(I + \Delta t \left(\frac{\partial f}{\partial x} \right)_{k-1} \right) S_{k-1} + \Delta t \left(\frac{\partial f}{\partial p} \right)_{k-1} \quad (\text{A3})$$

where I is a unit matrix. By expanding Eq. (A3) we get

$$\begin{aligned} S_k = & \left(\prod_{i=0}^{k-1} \left(I + \Delta t \left(\frac{\partial f}{\partial x} \right)_i \right) \right) S_0 + \Delta t \left(\prod_{i=1}^{k-1} \left(I + \Delta t \left(\frac{\partial f}{\partial x} \right)_i \right) \right) \left(\frac{\partial f}{\partial p} \right)_0 \\ & + \Delta t \left(\prod_{i=2}^{k-1} \left(I + \Delta t \left(\frac{\partial f}{\partial x} \right)_i \right) \right) \left(\frac{\partial f}{\partial p} \right)_1 + \dots + \Delta t \left(\frac{\partial f}{\partial p} \right)_{k-1} \end{aligned} \quad (\text{A4})$$

It can be reformulated as

$$\begin{aligned}
S_k &= \left(\prod_{i=0}^{k-1} \left(I + \Delta t \left(\frac{\partial f}{\partial x} \right)_i \right) \right) S_0 + W_0 \left(\frac{\partial f}{\partial p} \right)_0 + W_1 \left(\frac{\partial f}{\partial p} \right)_1 + \dots + W_{k-1} \left(\frac{\partial f}{\partial p} \right)_{k-1} \\
&= \left(\prod_{i=0}^{k-1} \left(I + \Delta t \left(\frac{\partial f}{\partial x} \right)_i \right) \right) S_0 + \sum_{j=0}^{k-1} W_j \left(\frac{\partial f}{\partial p} \right)_j
\end{aligned} \tag{A5}$$

Since $S_0 = \left(\frac{\partial x}{\partial p} \right)_0$ is the sensitivity at the initial state $x(t_0) = x_0$, there are two possible cases:

Case 1: $x(t_0) = x_0$ is a steady state. Then

$$S_0 = \left(\frac{\partial x}{\partial p} \right)_0 = \left(\left(\frac{\partial f}{\partial x} \right)_0 \right)^{-1} \left(\frac{\partial f}{\partial p} \right)_0 \tag{A6}$$

Case 2: $x(t_0) = x_0$ is not a steady state. Then we can consider that $x(t_0) = x_0$ is evolved from a steady state $x(-l) = x_{-l}$ at time point $t = -l$. According to Eq. (A5)

$$\begin{aligned}
S_0 &= \left(\prod_{i=-l}^{-1} \left(I + \Delta t \left(\frac{\partial f}{\partial x} \right)_i \right) \right) S_{-l} + \sum_{j=-l}^{-1} W_j \left(\frac{\partial f}{\partial p} \right)_j \\
&= \left(\prod_{i=-l}^{-1} \left(I + \Delta t \left(\frac{\partial f}{\partial x} \right)_i \right) \right) \left(\left(\frac{\partial f}{\partial x} \right)_{-l} \right)^{-1} \left(\frac{\partial f}{\partial p} \right)_{-l} + \sum_{j=-l}^{-1} W_j \left(\frac{\partial f}{\partial p} \right)_j \\
&= \tilde{W}_{-l} \left(\frac{\partial f}{\partial p} \right)_{-l} + \sum_{j=-l}^{-1} W_j \left(\frac{\partial f}{\partial p} \right)_j
\end{aligned} \tag{A8}$$

In both cases, S_0 has a linear relation with $\left(\frac{\partial f}{\partial p} \right)_j$. Then from Eq. (A5) there is

$$S_k = \left(\frac{\partial x}{\partial p} \right)_k = \sum_{j=0}^{k-1} V_j \Delta t \left(\frac{\partial f}{\partial p} \right)_j \tag{A9}$$

where V_j is a matrix computed at the discrete time point j . From Eq. (A9), for $\Delta t \rightarrow 0$, the sensitivity matrix can be expressed as

$$S = \int_{t_0}^t V(\tau) \left(\frac{\partial f}{\partial p} \right) d\tau \tag{A10}$$

2) The partial derivative functions of the three-step-pathway model

According to Eqs. (14-21) in the paper the functions to be partially derived are

$$f_1 = \frac{p_1}{1 + \left(\frac{P}{p_2}\right)^{p_3} + \left(\frac{p_4}{S}\right)^{p_5}} - p_6 x_1 \quad (\text{B1})$$

$$f_2 = \frac{p_7}{1 + \left(\frac{P}{p_8}\right)^{p_9} + \left(\frac{p_{10}}{x_7}\right)^{p_{11}}} - p_{12} x_2 \quad (\text{B2})$$

$$f_3 = \frac{p_{13}}{1 + \left(\frac{P}{p_{14}}\right)^{p_{15}} + \left(\frac{p_{16}}{x_8}\right)^{p_{17}}} - p_{18} x_3 \quad (\text{B3})$$

$$f_4 = \frac{p_{19} x_1}{p_{20} + x_1} - p_{21} x_4 \quad (\text{B4})$$

$$f_5 = \frac{p_{22} x_2}{p_{23} + x_2} - p_{24} x_5 \quad (\text{B5})$$

$$f_6 = \frac{p_{25} x_3}{p_{26} + x_3} - p_{27} x_6 \quad (\text{B6})$$

$$f_7 = \frac{p_{28} x_4 (S - x_7)}{p_{29} \left(1 + \frac{S}{p_{29}} + \frac{x_7}{p_{30}}\right)} - \frac{p_{31} x_5 (x_7 - x_8)}{p_{32} \left(1 + \frac{x_7}{p_{32}} + \frac{x_8}{p_{33}}\right)} \quad (\text{B7})$$

$$f_8 = \frac{p_{31} x_5 (x_7 - x_8)}{p_{32} \left(1 + \frac{x_7}{p_{32}} + \frac{x_8}{p_{33}}\right)} - \frac{p_{34} x_6 (x_8 - P)}{p_{35} \left(1 + \frac{x_8}{p_{35}} + \frac{P}{p_{36}}\right)} \quad (\text{B8})$$

From Eq. (B1),

$$\frac{\partial f_1}{\partial p_1} = \frac{1 + \left(\frac{P}{p_2}\right)^{p_3} + \left(\frac{p_4}{S}\right)^{p_5}}{\left[1 + \left(\frac{P}{p_2}\right)^{p_3} + \left(\frac{p_4}{S}\right)^{p_5}\right]^2} \quad (\text{B9})$$

$$\frac{\partial f_1}{\partial p_2} = \frac{\frac{p_1 p_3}{p_2} \left(\frac{P}{p_2}\right)^{p_3}}{\left[1 + \left(\frac{P}{p_2}\right)^{p_3} + \left(\frac{p_4}{S}\right)^{p_5}\right]^2} \quad (\text{B10})$$

$$\frac{\partial f_1}{\partial p_3} = \frac{-p_1 \left(\frac{P}{p_2}\right)^{p_3} \ln\left(\frac{P}{p_2}\right)}{\left[1 + \left(\frac{P}{p_2}\right)^{p_3} + \left(\frac{p_4}{S}\right)^{p_5}\right]^2} \quad (\text{B11})$$

$$\frac{\partial f_1}{\partial p_4} = \frac{-\frac{p_1 p_5}{p_4} \left(\frac{p_4}{S}\right)^{p_5}}{\left[1 + \left(\frac{P}{p_2}\right)^{p_3} + \left(\frac{p_4}{S}\right)^{p_5}\right]^2} \quad (\text{B12})$$

$$\frac{\partial f_1}{\partial p_5} = \frac{-p_1 \left(\frac{p_4}{S}\right)^{p_5} \ln\left(\frac{p_4}{S}\right)}{\left[1 + \left(\frac{P}{p_2}\right)^{p_3} + \left(\frac{p_4}{S}\right)^{p_5}\right]^2} \quad (\text{B13})$$

$$\frac{\partial f_1}{\partial p_6} = -x_1 \quad (\text{B14})$$

It can be clearly seen from Eqs. (B9-B13) that these partial derivative functions depend only on the parameters and controls. Thus $\frac{\partial f_1}{\partial p_1}, \frac{\partial f_1}{\partial p_2}, \dots, \frac{\partial f_1}{\partial p_5}$ are pairwise linearly dependent. From Eq. (B14),

$\frac{\partial f_1}{\partial p_6}$ depends on a state variable which will be a time-dependent profile and thus is linearly

independent with the other partial derivative functions.

From Eq. (B2),

$$\frac{\partial f_2}{\partial p_7} = \frac{1 + \left(\frac{P}{p_8}\right)^{p_9} + \left(\frac{p_{10}}{x_7}\right)^{p_{11}}}{\left[1 + \left(\frac{P}{p_8}\right)^{p_9} + \left(\frac{p_{10}}{x_7}\right)^{p_{11}}\right]^2} \quad (\text{B15})$$

$$\frac{\partial f_2}{\partial p_8} = \frac{\frac{p_7 p_9}{p_8} \left(\frac{P}{p_8}\right)^{p_9}}{\left[1 + \left(\frac{P}{p_8}\right)^{p_9} + \left(\frac{p_{10}}{x_7}\right)^{p_{11}}\right]^2} \quad (\text{B16})$$

$$\frac{\partial f_2}{\partial p_9} = \frac{-p_7 \left(\frac{P}{p_8}\right)^{p_9} \ln\left(\frac{P}{p_8}\right)}{\left[1 + \left(\frac{P}{p_8}\right)^{p_9} + \left(\frac{p_{10}}{x_7}\right)^{p_{11}}\right]^2} \quad (\text{B17})$$

$$\frac{\partial f_2}{\partial p_{10}} = \frac{-\frac{p_7 p_{11}}{p_{10}} \left(\frac{p_{10}}{x_7}\right)^{p_{11}}}{\left[1 + \left(\frac{P}{p_8}\right)^{p_9} + \left(\frac{p_{10}}{x_7}\right)^{p_{11}}\right]^2} \quad (\text{B18})$$

$$\frac{\partial f_2}{\partial p_{11}} = \frac{-p_7 \left(\frac{p_{10}}{x_7}\right)^{p_{11}} \ln\left(\frac{p_{10}}{x_7}\right)}{\left[1 + \left(\frac{P}{p_8}\right)^{p_9} + \left(\frac{p_{10}}{x_7}\right)^{p_{11}}\right]^2} \quad (\text{B19})$$

$$\frac{\partial f_2}{\partial p_{12}} = -x_2 \quad (\text{B20})$$

Based on Eqs. (B16-B17), we have

$$\frac{\partial f_2}{\partial p_8} = \left[\frac{-\frac{p_9}{p_8}}{\ln\left(\frac{P}{p_8}\right)} \right] \frac{\partial f_2}{\partial p_9} \quad (\text{B21})$$

Since the coefficient in Eq. (B21) only depends on parameters and a control variable P , $\frac{\partial f_2}{\partial p_8}$, $\frac{\partial f_2}{\partial p_9}$ are

linearly dependent. From Eqs. (B15-B18) it can be seen that

$$\frac{\partial f_2}{\partial p_7} - \left[\frac{1 + \left(\frac{P}{p_8}\right)^{p_9}}{\frac{p_7 p_9}{p_8} \left(\frac{P}{p_8}\right)^{p_9}} \right] \frac{\partial f_2}{\partial p_8} + \left(\frac{p_{10} p_{11}}{p_7} \right) \frac{\partial f_2}{\partial p_{10}} = 0 \quad (\text{B22})$$

$$\frac{\partial f_2}{\partial p_7} + \left[\frac{1 + \left(\frac{P}{p_8}\right)^{p_9}}{p_7 \left(\frac{P}{p_8}\right)^{p_9} \ln\left(\frac{P}{p_8}\right)} \right] \frac{\partial f_2}{\partial p_9} + \left(\frac{p_{10} p_{11}}{p_7} \right) \frac{\partial f_2}{\partial p_{10}} = 0 \quad (\text{B23})$$

Again, the coefficients in Eqs. (B22-B23) only depend on the parameters and the control variable P ,

therefore, two different groups, $\frac{\partial f_2}{\partial p_7}$, $\frac{\partial f_2}{\partial p_8}$, $\frac{\partial f_2}{\partial p_{10}}$ and $\frac{\partial f_2}{\partial p_7}$, $\frac{\partial f_2}{\partial p_9}$, $\frac{\partial f_2}{\partial p_{10}}$ are linearly dependent,

respectively. Similarly, according to Eqs. (B19-B20), $\frac{\partial f_2}{\partial p_{11}}$, $\frac{\partial f_2}{\partial p_{12}}$ are different from the other partial

derivative functions and thus linearly independent with each other and also with other partial derivative functions.

Similar results can be obtained by comparing the partial derivative functions of Eq. (B3), since Eq.

(B3) has the similar structure as Eq. (B2). Therefore, $\frac{\partial f_3}{\partial p_{14}}$, $\frac{\partial f_3}{\partial p_{15}}$ are linearly dependent in pair,

$\frac{\partial f_3}{\partial p_{13}}$, $\frac{\partial f_3}{\partial p_{14}}$, $\frac{\partial f_3}{\partial p_{16}}$ and $\frac{\partial f_3}{\partial p_{13}}$, $\frac{\partial f_3}{\partial p_{15}}$, $\frac{\partial f_3}{\partial p_{16}}$ are linearly dependent in two groups, respectively.

From Eq. (B4),

$$\frac{\partial f_4}{\partial p_{19}} = \frac{x_1}{p_{20} + x_1} \quad (\text{B24})$$

$$\frac{\partial f_4}{\partial p_{20}} = \frac{-p_{19}x_1}{(p_{20} + x_1)^2} \quad (\text{B25})$$

$$\frac{\partial f_4}{\partial p_{20}} = -x_4 \quad (\text{B26})$$

It can be clearly seen that $\frac{\partial f_4}{\partial p_{19}}, \frac{\partial f_4}{\partial p_{20}}, \frac{\partial f_4}{\partial p_{21}}$ are linearly independent. Similarly, according to Eqs.

(B5-B6), there are no linear dependences among $\frac{\partial f_5}{\partial p_{22}}, \frac{\partial f_5}{\partial p_{23}}, \frac{\partial f_5}{\partial p_{24}}$ and $\frac{\partial f_6}{\partial p_{25}}, \frac{\partial f_6}{\partial p_{26}}, \frac{\partial f_6}{\partial p_{27}}$.

From Eq. (B7),

$$\frac{\partial f_7}{\partial p_{28}} = \frac{\frac{x_4(S-x_7)}{p_{29}} \left(1 + \frac{S}{p_{29}} + \frac{x_7}{p_{30}} \right)}{\left(1 + \frac{S}{p_{29}} + \frac{x_7}{p_{30}} \right)^2} \quad (\text{B27})$$

$$\frac{\partial f_7}{\partial p_{29}} = \frac{-\frac{p_{28}x_4(S-x_7)}{p_{29}^2} \left(1 + \frac{x_7}{p_{30}} \right)}{\left(1 + \frac{S}{p_{29}} + \frac{x_7}{p_{30}} \right)^2} \quad (\text{B28})$$

$$\frac{\partial f_7}{\partial p_{30}} = \frac{\frac{p_{28}x_4(S-x_7)}{p_{29}p_{30}} \left(\frac{x_7}{p_{30}} \right)}{\left(1 + \frac{S}{p_{29}} + \frac{x_7}{p_{30}} \right)^2} \quad (\text{B29})$$

$$\frac{\partial f_7}{\partial p_{31}} = \frac{-\frac{x_5(x_7-x_8)}{p_{32}} \left(1 + \frac{x_7}{p_{32}} + \frac{x_8}{p_{33}} \right)}{\left(1 + \frac{x_7}{p_{32}} + \frac{x_8}{p_{33}} \right)^2} \quad (\text{B30})$$

$$\frac{\partial f_7}{\partial p_{32}} = \frac{\frac{p_{31}x_5(x_7-x_8)}{p_{32}^2} \left(1 + \frac{x_8}{p_{33}} \right)}{\left(1 + \frac{x_7}{p_{32}} + \frac{x_8}{p_{33}} \right)^2} \quad (\text{B31})$$

$$\frac{\partial f_7}{\partial p_{33}} = -\frac{p_{31}x_5(x_7 - x_8)\left(\frac{x_8}{p_{33}}\right)}{p_{32}p_{33}\left(1 + \frac{x_7}{p_{32}} + \frac{x_8}{p_{33}}\right)^2} \quad (\text{B32})$$

From Eqs. (B27-B29), $\frac{\partial f_7}{\partial p_{28}}, \frac{\partial f_7}{\partial p_{29}}, \frac{\partial f_7}{\partial p_{30}}$ are linearly dependent in one group. But $\frac{\partial f_7}{\partial p_{31}}, \frac{\partial f_7}{\partial p_{32}}, \frac{\partial f_7}{\partial p_{33}}$

are linearly independent, based on Eqs. (B30-B32).

From Eq. (B8),

$$\frac{\partial f_8}{\partial p_{34}} = \frac{-x_6(x_8 - P)\left(1 + \frac{x_8}{p_{35}} + \frac{P}{p_{36}}\right)}{p_{35}\left(1 + \frac{x_8}{p_{35}} + \frac{P}{p_{36}}\right)^2} \quad (\text{B33})$$

$$\frac{\partial f_8}{\partial p_{35}} = \frac{\frac{p_{34}x_6(x_8 - P)}{p_{35}^2}\left(1 + \frac{P}{p_{36}}\right)}{\left(1 + \frac{x_8}{p_{35}} + \frac{P}{p_{36}}\right)^2} \quad (\text{B34})$$

$$\frac{\partial f_8}{\partial p_{36}} = \frac{-\frac{p_{34}x_6(x_8 - P)}{p_{35}p_{36}}\left(\frac{P}{p_{36}}\right)}{\left(1 + \frac{x_8}{p_{35}} + \frac{P}{p_{36}}\right)^2} \quad (\text{B35})$$

It can be seen from Eqs. (B33-B35) that $\frac{\partial f_8}{\partial p_{35}}, \frac{\partial f_8}{\partial p_{36}}$ are linearly dependent, but $\frac{\partial f_8}{\partial p_{34}}$ is linearly

independent with $\frac{\partial f_8}{\partial p_{35}}, \frac{\partial f_8}{\partial p_{36}}$.

3) Table A1: P and S values for generating 5 datasets

Dataset	1	2	3	4	5
P	0.05000	0.36840	1.00000	0.09286	0.13572
S	10.0000	2.15440	0.10000	2.15440	2.15440